

PATENT ABSTRACTS OF JAPAN

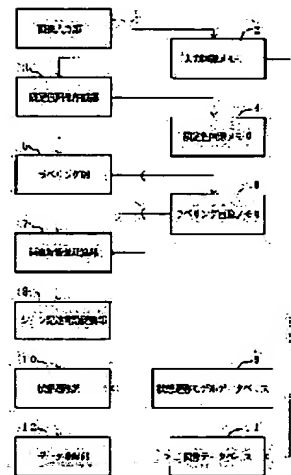
(11)Publication number : 08-249349  
(43)Date of publication of application : 27.09.1996

(51)Int.Cl. G06F 17/30

(21)Application number : 07-054794 (71)Applicant : SAKAUCHI MASAO  
SHARP CORP  
(22)Date of filing : 14.03.1995 (72)Inventor : SAKAUCHI MASAO  
ONO ATSUSHI

(54) IMAGE DATA BASE DEVICE

(57)Abstract:  
PURPOSE: To provide a versatile image data base device by enabling automatic extraction from an image by using a concept of higher level as a key word and lightening the operation burden at the time of registration, and enabling even an image which does not match with a state transition model to be registered and retrieved.  
CONSTITUTION: The image data base device divides an input image inputted by an image input part 1 into segments, converts an image feature quantity of each segment calculated by an image feature quantity calculation part 7 into a scene description language by a scene description language conversion part 8, shifts the scene description language of each segment to a concept of stratum as high as possible by a state transition part 10 on the basis of the state transition model, and registers the final state of each segment and the scene description language of a segment having no state transition as key words in an image data base 11 by a data registration part 12 together with the input image.



LEGAL STATUS  
[Date of request for examination] 15.12.1998  
[Date of sending the examiner's decision of rejection]  
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]  
[Date of final disposal for application]  
[Patent number] 3199976  
[Date of registration] 15.06.2001  
[Number of appeal against examiner's decision of rejection]  
[Date of requesting appeal against examiner's decision of rejection]  
[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

## \* NOTICES \*

**Japan Patent Office is not responsible for any damages caused by the use of this translation.**

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Industrial Application] this invention accumulates image information in large quantities, and relates to the image database equipment which carries out automatic extracting of the keyword from a picture especially in the case of registration of a picture about the image database equipment with which a picture can be searched by giving a keyword.

[0002]

[Description of the Prior Art] The expectation for the image database equipment for searching a required picture efficiently from the image information accumulated in large quantities in recent years is growing. Although the image database equipment with which the picture corresponding to this keyword is extracted is known if a registrant registers simultaneously the keyword which is the semantic information on the picture by manual input and a keyword is inputted by the reference person at the time of reference in case a picture is registered conventionally Since it is not efficient that a registrant inputs a keyword manually to all pictures, the various attempts for performing automatic extracting of a keyword from a picture are made using an image recognition or image comprehension technology.

[0003] As such an attempt, for example, the image database equipment with which (1) reference person inputs a sensibility word and the picture corresponding to the sensibility word is searched (refer to a "picture reference which was adapted for degree of subjective similar": Information Processing Society of Japan paper magazine, Vol.31, and no.2 (1990)), (2) The image database equipment (73 to "image database in consideration of keyword automatic extracting": Information Processing Society of Japan technical report '91-valve flow coefficient-1 reference) using the state-transition model etc. is known.

[0004]

[Problem(s) to be Solved by the Invention] However, the above-mentioned conventional composition (1) and (2) have the following troubles, respectively.

[0005] First, it is the composition of creating the map to the sensibility word which reflects a user's subjectivity scale from the color feature which a picture has based on a sample picture, and referring to the equipment of the above (1) by inputting this sensibility word at the time of reference. For this reason, while choosing and using a sample picture suitable at the time of map creation, since a sensibility word is subjective, it is lacking in versatility, and it has the trouble that adjustment of a map is required, by the reference person.

[0006] Moreover, it has the trouble that it is lacking in versatility since extraction of a keyword is completely dependent on a state-transition model, and it cannot refer to the equipment of the above (2) at all to what does not correspond to a model since a keyword is not given. Moreover, at the time of reference, it is required for a reference person to input only superordinate concepts, such as "soccer", as a keyword, scene description becomes a thing only depending on a model, and a reference person's will is not reflected.

[0007] this invention aims at realizing the high image database equipment of versatility while it was made in view of each above-mentioned trouble and can extract the keyword of more advanced conceptual level from a picture automatically at the time of registration of a picture.

[0008]

[Means for Solving the Problem] In order to solve the above-mentioned technical problem, the image database equipment of this invention according to claim 1 A picture storage means to match a picture with a keyword and to memorize it, and a picture input means to input a picture, The picture division means which carries out field division of the input picture, and a picture characteristic quantity calculation means to calculate the picture characteristic quantity of each divided field, A scene description language conversion means to change picture characteristic quantity into a scene description language, A state-transition model storage means to memorize a state-transition model, and a state-transition means to perform processing which makes the state of each field change to a high order hierarchy as long as the scene description language of each field suits the changes rule of a state-transition model, About the field where the state changed by processing of the above-mentioned state-transition means, make the final state of this field into a keyword, and a scene description language is made into a keyword about the field where a state did not change. It is characterized by having the registration means made to memorize to the above-mentioned picture storage means with the above-mentioned input picture.

[0009] Image database equipment according to claim 2 is set to equipment according to claim 1. A reference demand input means by which a reference person inputs a reference demand, and a reference demand analysis means to create the reference keyword at the time of analyzing the inputted reference demand and performing reference, While having further a picture

reference means to search a picture from a picture storage means according to the above-mentioned reference keyword, and a picture presentation means to show the picture of a reference result to a reference person. When the above-mentioned reference demand analysis means is in the state where the inputted reference demand exists in a state-transition model, while a changes rule is gone back from this state and this state and the above-mentioned picture reference means is passed by making into a reference keyword all the states that can reach. When the inputted reference demand does not exist as a state in a state-transition model, it is characterized by passing the above-mentioned picture reference means by making into a reference keyword the scene description language which the reference person was made to input a scene description language as a reference demand by the reference demand input means, and was inputted.

[0010] In equipment according to claim 1, image database equipment according to claim 3 is characterized by the above-mentioned registration means making it memorize to the above-mentioned picture storage means by making the above-mentioned positional information into a keyword about the field where the state changed by processing of a state-transition means in addition to the above-mentioned final state while a scene description language contains the positional information which shows the position of each field in an input picture.

[0011] Image database equipment according to claim 4 is set to equipment according to claim 2. In case the above-mentioned picture reference means searches a picture according to a reference keyword, when a reference keyword is in the state in a state-transition model. While giving a score to the picture which has a reference keyword and a keyword in agreement according to the hierarchy in the state-transition model of this state, when a reference keyword is a scene description language. The scene description language in which it moves and the state of the least-significant hierarchy of the state-transition model which can reach has the keyword of the picture memorized by the picture storage means to a changes rule. The scene description language as the above-mentioned reference keyword is compared, and while giving a score to the picture memorized by the above-mentioned picture storage means according to the degree of coincidence, the above-mentioned picture presentation means is characterized by showing the picture to which the score higher than a predetermined score was given as a reference result.

[0012]

[Function] According to composition according to claim 1, the picture characteristic quantity calculated from each field of the divided input picture is changed into a scene description language. As long as the scene description language of each field suits the changes rule of a state-transition model, the final state of the field about the field where the state of each field changed [ in ] to the superordinate concept, and the state changed as a keyword. Moreover, about the field where a state did not change, the scene description language of the field is memorized by the picture storage means with an input picture as a keyword.

[0013] Since a scene description language is automatically given as a keyword and is memorized by the picture storage means also to the picture which has by this the concept which does not suit a state-transition model. Since a registrant does not need to give a keyword in case a picture is registered, while becoming possible to mitigate a registrant's work burden, if a scene description language is given as a keyword, the above-mentioned picture can be searched at the time of reference. That is, flexible reference is attained, while the versatility of image database equipment improves, since the kind of picture in which storage and reference are possible is not limited with the state-transition model which image database equipment holds.

[0014] When the reference demand which the reference person inputted exists as a state in the state-transition model which image database equipment holds according to composition according to claim 2. A changes rule is gone back from this state and the above-mentioned state. all the states of the low rank hierarchy which can reach. A picture reference means is passed as a reference keyword, the picture in which a picture reference means has this reference keyword and a keyword in agreement is searched from a picture storage means, and the reference result is shown to a reference person. On the other hand, when the reference demand which the reference person inputted does not exist as a state in the state-transition model which image database equipment holds, the scene description language which the reference person was made to input by the reference input means is passed to a picture reference means as a reference keyword, and a picture reference means searches a picture according to the scene description language passed as a reference keyword.

[0015] Even if it is the case where reference of the picture which does not suit by this the state-transition model which image database equipment holds is required, a scene description language is made to input as a reference demand, and it becomes possible to search a picture according to this scene description language. Consequently, it is lost with image database equipment that the kind of picture in which storage and reference are possible is limited with a state-transition model, and it becomes possible to raise the versatility of image database equipment.

[0016] According to composition according to claim 3, about the field where the state changed to the high order hierarchy by processing of a state-transition means. Since the position in the input picture of the field is registered [ final state / of a field ] with an input picture as a keyword. In case it searches, when a reference person performs the specification a "person" as a state of a "center section" and a field as positional information, it becomes possible to give the conditions of "the picture to which a person exists in a center section", and to search a picture. Consequently, retrieval effectiveness can be raised, while becoming possible to search by pointing to the picture image for [ which the reference person is holding ] reference to image database equipment more concretely and attaining the flexible reference in which a reference person's volition was reflected.

[0017] According to composition according to claim 4, the picture searched as a reference keyword the state in a state-transition model. The picture searched considering the scene description language as a reference keyword by giving such a high score that a hierarchy becoming high according to the hierarchy in the state-transition model of the above-mentioned state. The scene description language which goes back a changes rule from the keyword of the picture memorized by the picture storage means, and the state of the least-significant hierarchy of the state-transition model which can reach has, Only the picture to which the scene description language as the above-mentioned reference keyword was compared, such a high score that the degree of

coincidence becomes high was given to, and the score higher than a predetermined score was given is shown as a reference result to a reference person.

[0018] Since a picture with high possibility of being the picture for which a reference person asks will be shown by this, while becoming possible to raise the retrieval effectiveness at the time of a reference person choosing a desired picture from the picture shown as a reference result etc., possibility that the picture which was widely different from the picture for which a reference person asks will be shown can be decreased, and the reliability of image database equipment can be raised.

[0019]

[Example] It will be as follows if one example of this invention is explained based on drawing 1 or drawing 6. As the image database equipment in this example is image database equipment with which a color picture is registered and searched and is shown in drawing 1 as composition about registration of a picture The image database 11 which accumulates a picture and its keyword, and the picture input section 1 (picture input means) for inputting the picture registered into a database, The input image memory 2 for accumulating the picture inputted from this picture input section 1, The limited color picture creation section 3 (picture division means) which divides the whole picture into two or more fields, and creates a limited color picture while inputting the full color picture accumulated at this input image memory 2 and carrying out subtractive color of the color number, The limited color picture memory 4 which accumulates this limited color picture, and the labeling section 5 which carries out labeling of the above-mentioned limited color picture, and creates a labeling picture, The labeling picture memory 6 which accumulates a labeling picture, and the picture characteristic quantity calculation section 7 (picture characteristic quantity calculation means) which calculates the picture characteristic quantity of each field, The scene description language transducer 8 (scene description language conversion means) which changes picture characteristic quantity into a scene description language, The state-transition model database 9 (state-transition model storage means) which accumulated the rule for the state of each above-mentioned field changing, It has the state-transition section 10 (state-transition means) which makes the state of each field change, referring to this state-transition model database 9, and the data registration section 12 (registration means) which registers the input picture of the input image memory 2 to the above-mentioned image database 11 with a keyword.

[0020] Here, a state-transition model is explained. It is a model for bringing more the state of the segment which serves as a state-transition model from a state transition diagram as shown in drawing 6, and a changes rule, and mentions later operation in which a state will change to a high order hierarchy as an arrow shows if the changes rule in this drawing set [ in / states, such as "Human" and "Human-hair", / for example ] up beforehand is filled by repeating as long as the changes rule filled exists close to a superordinate concept.

[0021] Here, registration of the color picture by the above-mentioned composition is explained below, referring to the flow chart shown in drawing 2. In addition, in this example, the color picture registered presupposes that each element of (Red R), (Green G), and blue (B) is the full color picture expressed by the 8 bits (256 gradation from 0 to 255) digital data, respectively.

[0022] First, the color picture which it is going to register into an image database is inputted by the picture input section-1, and is accumulated at the input image memory 2 (it writes like S1 Step 1 and the following).

[0023] Next, although a body etc. is recognized from this input picture, it is necessary to cut down first the body which serves as a candidate for recognition from a picture for this recognition. Here, since an input picture is a full color picture, it uses the technique of the color segmentation which extracts the field constituted by the same color. That is, create the limited color picture which limited the color to about 16 colors from the full color input picture which the limited color picture creation section 3 took out from the input image memory 2 (S2), and the labeling section 5 sets in the limited color picture created by the above S2. The same label is given to the connection field of the same color, and a picture is divided into two or more fields by making into one connection field the field which has the same label (S3). Henceforth, suppose that each of these divided fields is called segment.

[0024] Next, the picture characteristic quantity calculation section 7 computes picture characteristic quantity, such as the pulse duty factor and circularity to area, a circumscription rectangle, a color, the direction of a main shaft, and the direction of a main shaft, and the center of gravity, about each segment, respectively, and (S4) and the scene description language transducer 8 create the scene description language of each segment from the computed picture characteristic quantity (S5). in addition, this scene description language is in every direction in the size of a segment, and the whole input picture -- when it trichotomizes, respectively and divides into nine blocks, it consists of two or more kinds of parameters which express the block position where an applicable segment is contained, the color of a segment, the configuration of a segment, and the direction of a main shaft of a segment, respectively

[0025] In S6 continuing, the state-transition section 10 does as follows the work which determines the concept of each segment based on the scene description language of each segment, and the state-transition model accumulated at the state-transition model database 9.

[0026] That is, the state-transition section 10 will make the state of this segment change to the hierarchy on one, if it investigates whether the scene description language created by the above S5 to each segment is filling the changes rule of a state-transition model sequentially from the subordinate concept of a state-transition model and the changes rule is filled.

[0027] A state-transition model as shown in drawing 6 is more specifically given, and if the parameter which shows the color of the segment in the scene description language of a certain segment is "black", this segment will fill the changes rule, will change to 1 hierarchy high order from "Color-segment" which is the concept of the least significant, and will serve as "Black-segment." Furthermore, when the scene description language of this segment fulfills the conditions which may be the picture of human being's hair portion, it means that the changes rule was filled further and changes to 1 hierarchy high order to a pan. Thereby, this segment is recognized as a segment which has conceptual "Human-hair" that it is human being's hair.

[0028] To the above-mentioned state-transition model, moreover, for the excessive rate at the time of the field division in S3, shading, etc. Since what is originally one body corresponds to the body divided into two or more segments, the body which consists of two or more colors Comparatively, the changes rule as a field integrated rule which performs only integration of segments, without making a state change in the hierarchy of a high order is also contained, the segments which fill this field integrated rule are unified, and they are recognized as a segment which has one new concept.

[0029] As a segment which a state-transition model as shown in drawing 6 is more specifically given, and is contained in an input picture When the segment which has above-mentioned conceptual "Human-hair", and the segment which has conceptual "Human-skin" that it is the picture of human being's skin portion exist These segments are unified as what fills a field integrated rule, and are recognized as one segment which has conceptual "Human-face" that it is the picture of human being's face portion.

[0030] After it is repeated until no state of the segments in an input picture stops changing, and this operation is completed, the data registration section 12 registers operation of Sabove-mentioned 6 into an image database 11 with the picture inputted by the above S1 by making into a keyword concepts, such as the final state called for by the above S6 from each segment, i.e., above-mentioned "Human-face" etc., (S7). In addition, about the segment which changed to the concept of a high order rather than the predetermined hierarchy, the positional parameter of a scene description language is also registered with the above-mentioned keyword to an image database 11 at this time. Moreover, about the segment to which a state did not change at all in the above S6, a scene description language is registered as a keyword.

[0031] In addition, with the above-mentioned state-transition model, on the low rank hierarchy of a state-transition model, the importance of a scene description language becomes low, and the field integrated rule is set up so that segments may be unified preferentially as importance is attached to a scene description language, the changes rule is set up so that a state may be made to change, and it becomes a high order hierarchy. Thus, on the low rank hierarchy, while raising versatility by making a state change by the scene description language, to the high order hierarchy, the reliability of recognition is raised by seldom depending on a scene description language, but seeing the relation between segments etc. globally, and performing a state transition.

[0032] In addition, here explains creation of the limited color picture in the above S2 in more detail based on the flow chart shown in drawing 3 .

[0033] First, the limited color picture creation section 3 creates the histogram of a full color input picture (S21). Since each of R-G-B is expressed by 256 gradation, each color in a picture is expressed with this example as a point of a RGB sky throughout which is among the cube of 256 (from 0 to 255) one side. It divides into the cube of the this 1 called cell in the cube of 256 one side side 8. Next, the RGB value of the pixel in a picture is calculated, it asks for to which cell the pixel belongs all over RGB space, and only 1 makes the frequency value of the cell to which the pixel belongs increase. This processing is performed to all pixels and a histogram is created.

[0034] Next, the mode cell which is a cell with the largest frequency value is detected in the created histogram (S22); the color which this mode cell expresses is made into a representation color, and the color of the cell contained in the sphere of the radius to which this mode cell is beforehand given as a center is replaced in the above-mentioned representation color. Thus, the cell as which the color was determined is excepted from an object, and it returns to S22, and the same processing is repeated until an object cell is lost. Subtractive color of the color number of a full color picture can be carried out to about 16 colors by making the above-mentioned radius at the time of performing these processings adjust suitably. That is, when there is little color number, the above-mentioned radius can be made small, the number of the cells which can be expressed with one representation color by enlarging the above-mentioned radius in many can be adjusted, and the color number can be brought close to the number of hope.

[0035] Next, it explains, referring to drawing 4 and drawing 5 about reference operation of the picture by the composition with which the image database equipment of this example is equipped, and this composition, in order that a reference person may search a picture from the image database 11 into which the picture is registered with the keyword, as explained above.

[0036] Further in addition to the composition about registration of the picture mentioned above, as shown in drawing 4 , the image database equipment of this example The reference demand input section 21 (reference demand input means) for the reference person who searches inputting a reference demand, The reference demand analysis section 22 (reference demand analysis means) which analyzes the inputted reference demand, It is the composition equipped with the reference section 23 (picture reference means) which searches the picture which corresponds based on the analysis result of a reference demand from an image database 11, and the picture presentation section 24 (picture presentation means) which shows a reference person the picture searched by the reference section 23.

[0037] In the above-mentioned composition, in searching the picture accumulated at the image database 11, a reference person inputs first the name of the structure which constitutes the picture to search by the reference demand input section 21 as a reference demand (S11). When it explains more concretely and a reference person, for example, wants to search a picture including the person, it is "Human" as a name of a structure. It will input.

[0038] If a reference demand is inputted, the reference demand analysis section 22 will look for the state-transition model database 9. The concept which is in agreement with the name of the structure inputted as a reference demand by S11 If contained in the state-transition model accumulated at the state-transition model database 9 (S12), it will return to S11 and the reference demand input section 21 will direct to input the reference demand about the position of the above-mentioned structure in a picture to a reference person.

[0039] Thereby, a reference person becomes possible [ creating a reference demand, such as "a picture in which a person is right in the middle", ]. Consequently, while becoming possible to search by pointing to the picture image for [ which the reference

person is holding ] reference to image database equipment more concretely and attaining the flexible reference reflecting a reference person's will, the effect that retrieval effectiveness improves is done so. However, a reference person does not necessarily need to follow the above-mentioned directions, and can also progress to S13 which is the following step, without inputting the reference demand about a position.

[0040] If the concept which is in agreement with the name of the structure inputted as a reference demand on the other hand does not exist in the state-transition model accumulated at the state-transition model database 9 (S12), it returns to S11 and the reference demand input section 21 makes a reference person input the reference demand by the scene description language. In addition, as described above as the scene description language, it is the position to a size and the whole input picture, the color, the configuration, and the direction of a main shaft of an object. However, it is not necessary to input these [ all ] as a reference demand.

[0041] Thus, when the position or a desired concept does not exist in a state-transition model if needed [ the concept and if needed ] which exist in a state-transition model, a scene description language is passed to the reference section 23 as a reference demand. In addition, the above S11 and S12 is repeated until a reference person ends creation of all reference demands (S13).

[0042] In addition, the user interface in the reference demand input section 21 for a reference person inputting a reference demand For example, if the concept included in the state-transition model in the menu screen by GUI (Graphical User Interface) is displayed as an icon and a reference person chooses a desired thing from these icons According to [ the concept corresponding to the selected icon is able to be passed to the reference section 23 as a reference demand, and ] this, the operability for a reference person can be raised. Moreover, when the concept corresponding to the structure of a request of a reference person is not displayed as an icon (i.e., when the desired concept is not included in a state-transition model), it can also be directed to make various kinds of parameters of a scene description language input in the pictures.

[0043] The reference section 23 searches the picture which corresponds from an image database 11 based on the above-mentioned reference demand (S14), and the picture presentation section 24 presents the reference result of the reference section 23 to a reference person (S15).

[0044] Here, the above-mentioned reference operation is explained more to a detail using the example of the state-transition model shown in drawing 6.

[0045] The reference demand analysis section 22 performs processing which is different whether the inputted reference demand is the concept which exists in a state-transition model, or it is a scene description language. Processing when the concept which exists in a state-transition model is first inputted as a reference demand is explained. Here, it is name "Human" of a structure by the reference person. As it is inputted as a reference demand and shown in drawing 6, it is "Human" to a state-transition model. Suppose that it exists as a concept.

[0046] When existing in a state-transition model judges the above-mentioned reference demand in the analysis of the reference demand analysis section 22, the reference demand analysis section 22 is this reference demand "Human". It considers as a reference keyword, and passes to the reference section 23, and the reference section 23 searches an image database 11 based on this reference keyword. If the picture whose "Human" is a keyword is registered into the image database 11 as a result of search, let this picture be a candidate picture. furthermore, the reference demand analysis section 22 -- the above-mentioned state-transition model -- setting -- "Human" from -- when the picture which has the keyword which corresponds with this reference keyword searches and exists [ whether it registers with the image database 11, and ], the score according to the hierarchy of this "Human-face" gives [ pass / the reference section 23 / by making into a reference keyword "Human-face" which is the concept which got down only from one hierarchy to the low rank ] by making this picture into a candidate picture In addition, the above-mentioned score is given according to the hierarchy of a state-transition model, and a high score is given, so that the hierarchy of a high order is approached.

[0047] After performing grant of the search and the score by the above reference keywords one by one to the hierarchy of the least significant of a state-transition model, the picture which has a bigger score than the threshold beforehand set up among the candidate picture and the picture which became is presented by the picture presentation section 24 as a picture of a reference result to a reference person. Moreover, when there are two or more kinds of reference demands, the picture to which the average of the score to all reference demands became larger than the above-mentioned threshold among the candidate picture and the picture which became is shown as a reference result.

[0048] Next, processing when the inputted reference demand is a scene description language is explained. Here, suppose that the state-transition model shown in drawing 6 is used.

[0049] In this case, the reference demand analysis section 22 is passed to the reference section 23 by making the inputted scene description language into a reference keyword. Next, what is registered into the image database 11 as a keyword of a registration picture among the concepts in the state-transition model of the state-transition model database 9 is chosen from a state-transition model. That is, if only "Human-face" shall be registered into the image database 11 with the picture as a keyword among the concepts which exist in the state-transition model shown, for example in drawing 6, this "Human-face" will be chosen.

[0050] And the reference section 23 performs comparison with the scene description language given as a reference demand in the concept of the least-significant hierarchy which reaches by going back a changes rule from "Human-face" to a low rank hierarchy. The changes rule in the low rank hierarchy of a state-transition model in more detail Each parameter of the scene description language which is given as power conditions which each parameter of a scene description language fulfills, and was given as a reference demand, The score as a candidate picture is given to the picture by which each parameter of the scene description language as the above-mentioned conditions is registered into the image database 11 by making above-mentioned "Human-face" into a keyword according to whether to be how much in agreement. In addition, the score given shall also become



high, so that the degree of coincidence of each above-mentioned parameter is high.

[0051] As mentioned above, after performing grant of the comparison based on each parameter of a scene description language, and a score to all the keywords registered into the image database 11, the picture which has a bigger score than the threshold beforehand set up among the candidate picture and the picture which became is presented by the picture presentation section 24 as a picture of a reference result to a reference person. Moreover, when there are two or more kinds of reference demands, the picture with the larger average of the score to all reference demands among a candidate picture and the picture which became than the above-mentioned threshold is shown as a reference result.

[0052] As mentioned above, the image database equipment of this example divides into two or more segments the full color picture which the picture input section inputted based on a color, and creates a scene description language from the picture characteristic quantity computed from each segment. Furthermore, in case the state of each segment is made to change to the concept of a high level as much as possible and a picture is registered into an image database by applying a state-transition model to this scene description language, the final state of each segment is registered as a keyword with a picture. Moreover, it is the composition of also registering the position of this segment in a picture with the above-mentioned final state about the segment which changed to the high order hierarchy, and registering a scene description language about the segment to which the changes to a high order hierarchy were not performed at this time.

[0053] Since an objective keyword is given \*\* [ according to / a registrant's subjectivity ] while the work burden of the registrant at the time of creation of an image database is sharply reducible, since the work which a registrant gives a keyword becomes unnecessary by this to each of the picture to register, the effect that the versatility and reliability of an image database can be raised is done so.

[0054] moreover, when searching the picture registered by doing in this way and the concept which exists in a state-transition model is given as a reference demand It searches for whether while searching an image database by making this concept into a keyword, the picture which made the keyword the subordinate concept in the changes rule which reaches to the above-mentioned concept is registered into the image database one by one to a least-significant concept. When the picture is registered, the score according to the hierarchy in the state-transition model of the concept as a keyword of the picture is given to the above-mentioned picture, and has composition shown to a reference person by making only a picture with this larger score than a predetermined threshold into a reference result.

[0055] Even when the concept as a reference demand which the reference person gave is not registered as a keyword by this, it is possible to raise possibility that the picture in which a reference person asks for the concept of the low rank hierarchy of the concept since it will be shown as a reference result if there is a picture registered as a keyword will be shown as a reference result. Consequently, the effect that the flexible reference based on the more abstract advanced concept is attained is done so. Moreover, it prevents that the picture for which a reference person asks, and the picture which was widely different are shown since only a picture with a larger score than a predetermined threshold is shown, and while a reference person becomes possible mitigating the burden which chooses a desired picture ], the reliability of an image database can be raised.

[0056] moreover, about the segment to which a state did not change with the state-transition model which this image database equipment holds in the case of registration of a picture A scene description language is registered with a picture as a keyword. in the case of reference When the concept which does not exist in the state-transition model which this image database equipment holds is given by the reference person as a reference demand While giving a reference person directions so that a reference demand on scene description language level may be inputted The changes rule in the concept of the least-significant hierarchy which can reach by going back a changes rule from the concept registered into the image database as a keyword, It is the composition that compare each parameter with the scene description language inputted by the reference person as a reference demand, give a score to the picture registered corresponding to the above-mentioned keyword according to the degree of coincidence, and this score presents only a larger picture than a predetermined threshold as a reference result.

[0057] The effect that an image database can raise the versatility of an image database only depending on a state-transition model also about the picture which has the concept which does not exist in the state-transition model which this image database equipment holds by this since it becomes possible to perform registration and reference by using a scene description language is done so.

[0058] In addition, although this invention is not limited to the above-mentioned example and this example explained the kind of parameter of a scene description language as the size, the position, the color, the configuration, and the direction of a main shaft of a segment, it is also possible to make other elements into a scene description language according to the kind of picture registered into a database. Moreover, according to the use purpose of the kind of picture, or a database to register, various models are applicable also about a state-transition model.

[0059]

[Effect of the Invention] As mentioned above, the image database equipment of this invention according to claim 1 A picture storage means to match a picture with a keyword and to memorize it, and a picture input means to input a picture, The picture division means which carries out field division of the input picture, and a picture characteristic quantity calculation means to calculate the picture characteristic quantity of each divided field, A scene description language conversion means to change picture characteristic quantity into a scene description language, A state-transition model storage means to memorize a state-transition model, and a state-transition means to perform processing which makes the state of each field change to a high order hierarchy as long as the scene description language of each field suits the changes rule of a state-transition model, It is composition equipped with the registration means which makes the final state of this field a keyword about the field where the state changed by processing of the above-mentioned state-transition means, and is made to memorize to the above-mentioned

picture storage means with the above-mentioned input picture by making a scene description language into a keyword about the field where a state did not change.

[0060] Since the kind of picture in which storage and reference are possible is not limited with the state-transition model which image database equipment holds while becoming possible to mitigate a registrant's work burden, since a scene description language is automatically given as a keyword also to the picture which has by this the concept which does not suit a state-transition model, the effect that the versatility of image database equipment improves does so.

[0061] A reference demand input means by which, as for image database equipment according to claim 2, a reference person inputs a reference demand, A reference demand analysis means to create the reference keyword at the time of analyzing the inputted reference demand and performing reference, While having further a picture reference means to search a picture from a picture storage means according to the above-mentioned reference keyword, and a picture presentation means to show the picture of a reference result to a reference person When the above-mentioned reference demand analysis means is in the state where the inputted reference demand exists in a state-transition model While a changes rule is gone back from this state and this state and the above-mentioned picture reference means is passed by making into a reference keyword all the states that can reach When the inputted reference demand does not exist as a state in a state-transition model, it is the composition passed to the above-mentioned picture reference means by making into a reference keyword the scene description language which the reference person was made to input a scene description language as a reference demand by the reference demand input means, and was inputted.

[0062] Even if it is the case where reference of the picture which does not suit by this the state-transition model which image database equipment holds is required, it becomes possible to make a scene description language input as a reference demand, and to search a picture. Consequently, while it is lost with image database equipment that the kind of picture in which storage and reference are possible is limited with a state-transition model and more flexible reference is attained, the effect of becoming possible to raise the versatility of image database equipment is done so.

[0063] Image database equipment according to claim 3 is composition which the above-mentioned registration means makes memorize to the above-mentioned picture storage means by making the above-mentioned positional information into a keyword about the field where the state changed by processing of a state-transition means in addition to the above-mentioned final state while a scene description language contains the positional information which shows the position of each field in an input picture.

[0064] Since this is enabled to input the body contained in a picture as a reference demand with the position, it becomes possible to search by pointing to the picture image for [ which the reference person is holding ] reference to image database equipment more concretely. Consequently, it depends reflecting a reference person's will, flexible reference is attained, and the effect that retrieval effectiveness can be raised is done so.

[0065] In case the above-mentioned picture reference means searches a picture according to a reference keyword, when a reference keyword is in the state in a state-transition model, image database equipment according to claim 4 While giving a score to the picture which has a reference keyword and a keyword in agreement according to the hierarchy in the state-transition model of this state, when a reference keyword is a scene description language The scene description language in which it moves and the state of the least-significant hierarchy of the state-transition model which can reach has the keyword of the picture memorized by the picture storage means to a changes rule, While giving a score to the picture which compares the scene description language as the above-mentioned reference keyword, and is memorized by the above-mentioned picture storage means according to the degree of coincidence, the above-mentioned picture presentation means is the composition of showing the picture to which the score higher than a predetermined score was given as a reference result.

[0066] Since a picture with high possibility of being the picture for which a reference person asks is shown by this While becoming possible to raise the retrieval effectiveness at the time of a reference person choosing a desired picture from the picture shown as a reference result etc. Since possibility that the picture which was widely different from the picture for which a reference person asks will be shown decreases, while a reference person can mitigate the burden which chooses a desired picture, the effect that the reliability of image database equipment can be raised is done so.

---

[Translation done.]



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平8-249349

(43)公開日 平成8年(1996)9月27日

(51)Int.Cl.<sup>6</sup>

G 0 6 F 17/30

識別記号

庁内整理番号

9194-5L

9194-5L

F I

G 0 6 F 15/40

15/401

技術表示箇所

3 7 0 B

3 1 0 A

審査請求 未請求 請求項の数4 O L (全 11 頁)

(21)出願番号 特願平7-54794

(22)出願日 平成7年(1995)3月14日

(71)出願人 595037261

坂内 正夫

神奈川県横浜市青葉区美しが丘2-56-7

(71)出願人 000005049

シャープ株式会社

大阪府大阪市阿倍野区長池町22番22号

(72)発明者 坂内 正夫

神奈川県横浜市青葉区美しが丘2-56-7

(72)発明者 小野 教史

大阪府大阪市阿倍野区長池町22番22号

シャープ株式会社内

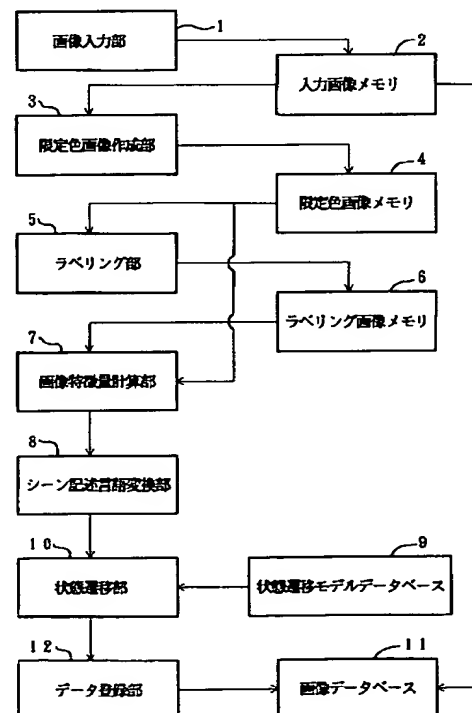
(74)代理人 弁理士 原 謙三

## (54)【発明の名称】 画像データベース装置

## (57)【要約】

【構成】 画像データベース装置は、画像入力部1が入力した入力画像をセグメントに分割し、画像特徴量計算部7により算出された各セグメントの画像特徴量をシーン記述言語変換部8がシーン記述言語に変換し、各セグメントのシーン記述言語を状態遷移部10が状態遷移モデルに基づいて可能な限り上位階層の概念へ遷移させ、データ登録部12が、各セグメントの最終状態、あるいは状態が遷移しなかったセグメントについてはシーン記述言語を、キーワードとして入力画像と共に画像データベース11へ登録する。

【効果】 より高度な概念をキーワードとして画像から自動抽出することができ、登録時の作業負担が軽減されると共に、状態遷移モデルに適合しない画像であっても登録・検索することが可能となり、汎用性の高い画像データベース装置が実現される。



## 【特許請求の範囲】

【請求項 1】画像をキーワードと対応づけて記憶する画像記憶手段と、画像を入力する画像入力手段と、入力画像を領域分割する画像分割手段と、分割された各領域の画像特徴量を求める画像特徴量算出手段と、画像特徴量をシーン記述言語に変換するシーン記述言語変換手段と、状態遷移モデルを記憶する状態遷移モデル記憶手段と、各領域の状態を上位階層へ遷移させる処理を、各領域のシーン記述言語が状態遷移モデルの遷移ルールに適合する限り行う状態遷移手段と、上記状態遷移手段の処理によって状態が遷移した領域については該領域の最終状態をキーワードとし、状態が遷移しなかった領域についてはシーン記述言語をキーワードとして、上記入力画像と共に上記画像記憶手段へ記憶させる登録手段とを備えていることを特徴とする画像データベース装置。

【請求項 2】検索者が検索要求を入力する検索要求入力手段と、入力された検索要求を解析して検索を実行する際の検索キーワードを作成する検索要求解析手段と、上記検索キーワードに従って画像記憶手段から画像を検索する画像検索手段と、検索結果の画像を検索者へ提示する画像提示手段とをさらに備えると共に、上記検索要求解析手段が、入力された検索要求が状態遷移モデル中に存在する状態である場合には、該状態および該状態から遷移ルールを逆行して到達可能な状態すべてを検索キーワードとして上記画像検索手段へ渡す一方、入力された検索要求が状態遷移モデル中に状態として存在しない場合には、検索者に検索要求入力手段によりシーン記述言語を検索要求として入力させ、入力されたシーン記述言語を検索キーワードとして上記画像検索手段へ渡すことを特徴とする請求項 1 記載の画像データベース装置。

【請求項 3】シーン記述言語が、入力画像における各領域の位置を示す位置情報を含むと共に、上記登録手段が、状態遷移手段の処理によって状態が遷移した領域については、上記最終状態に加えて上記位置情報をキーワードとして上記画像記憶手段へ記憶させることを特徴とする請求項 1 記載の画像データベース装置。

【請求項 4】上記画像検索手段が検索キーワードに従って画像を検索する際に、検索キーワードが状態遷移モデル中の状態である場合には、該状態の状態遷移モデルにおける階層に応じて、検索キーワードと一致するキーワードを有する画像に得点を付与する一方、検索キーワードがシーン記述言語である場合には、画像記憶手段に記憶されている画像のキーワードから遷移ルールを逆行して到達可能な状態遷移モデルの最下位階層の状態の持つシーン記述言語と、上記検索キーワードとしてのシーン記述言語とを比較し、その一致度に応じて上記の画像記憶手段に記憶されている画像に得点を付与すると共に、上記画像提示手段が、所定の得点よりも高い得点を付与された画像を検索結果として提示することを特徴とする

請求項 2 記載の画像データベース装置。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、画像情報を大量に蓄積し、キーワードを与えることによって画像の検索を行うことのできる画像データベース装置に関し、特に、画像の登録の際に画像からキーワードを自動抽出する画像データベース装置に関する。

## 【0002】

- 10 【従来の技術】近年、大量に蓄積された画像情報から必要な画像を効率良く検索するための画像データベース装置への期待が高まっている。従来、画像を登録する際に、登録者が、その画像の意味的な情報であるキーワードを手動入力によって同時に登録し、検索時に検索者によってキーワードが入力されるとこのキーワードに対応する画像が抽出される画像データベース装置が知られているが、すべての画像に対して登録者がキーワードを手動入力するのは効率的でないため、画像認識や画像理解技術を用いて、画像からキーワードの自動抽出を行うための種々の試みがなされている。

20 【0003】このような試みとして、例えば、(1) 検索者が感性語を入力し、その感性語に対応する画像を検索する画像データベース装置（「主観的類似度に応じた画像検索」：情報処理学会論文誌，Vol. 31, no. 2 (1990) 参照）、(2) 状態遷移モデルを用いた画像データベース装置（「キーワード自動抽出を考慮した画像データベース」：情報処理学会技報 '91-CV-73-1 参照）等が知られている。

## 【0004】

- 30 【発明が解決しようとする課題】しかしながら、上記した従来の構成(1)および(2)は、下記のような問題点をそれぞれ有している。

【0005】まず、上記(1)の装置では、画像の持つ色特徴から利用者の主観尺度を反映する感性語への写像を、サンプル画像に基づいて作成し、検索時には、この感性語を入力することにより検索を行う構成である。このため、写像作成時に適切なサンプル画像を選択して用いる必要があると共に、感性語は主観的なものであるので汎用性に乏しく、検索者によって写像の調整が必要であるという問題点を有している。

40 【0006】また、上記(2)の装置では、キーワードの抽出が状態遷移モデルに完全に依存しているために汎用性に乏しく、モデルに該当しないものに対しては、キーワードが付与されないために全く検索を行うことができないという問題点を有している。また、検索時には、検索者が例えば「サッカー」等の上位概念のみをキーワードとして入力することが必要であり、シーン記述はモデルのみに依存するものとなり、検索者の意志が反映されない。

- 50 【0007】本発明は上記した各問題点に鑑みなされた

## 3

もので、画像の登録時に、より高度な概念レベルのキーワードを画像から自動的に抽出できると共に、汎用性の高い画像データベース装置を実現することを目的としている。

## 【0008】

【課題を解決するための手段】上記の課題を解決するために、本発明の請求項1記載の画像データベース装置は、画像をキーワードと対応づけて記憶する画像記憶手段と、画像を入力する画像入力手段と、入力画像を領域分割する画像分割手段と、分割された各領域の画像特徴量を求める画像特徴量算出手段と、画像特徴量をシーン記述言語に変換するシーン記述言語変換手段と、状態遷移モデルを記憶する状態遷移モデル記憶手段と、各領域の状態を上位階層へ遷移させる処理を、各領域のシーン記述言語が状態遷移モデルの遷移ルールに適合する限り行う状態遷移手段と、上記状態遷移手段の処理によって状態が遷移した領域については該領域の最終状態をキーワードとし、状態が遷移しなかった領域についてはシーン記述言語をキーワードとして、上記入力画像と共に上記画像記憶手段へ記憶させる登録手段とを備えていることを特徴としている。

【0009】請求項2記載の画像データベース装置は、請求項1記載の装置において、検索者が検索要求を入力する検索要求入力手段と、入力された検索要求を解析して検索を実行する際の検索キーワードを作成する検索要求解析手段と、上記検索キーワードに従って画像記憶手段から画像を検索する画像検索手段と、検索結果の画像を検索者へ提示する画像提示手段とをさらに備えると共に、上記検索要求解析手段が、入力された検索要求が状態遷移モデル中に存在する状態である場合には、該状態および該状態から遷移ルールを逆行して到達可能な状態すべてを検索キーワードとして上記画像検索手段へ渡す一方、入力された検索要求が状態遷移モデル中に状態として存在しない場合には、検索者に検索要求入力手段によりシーン記述言語を検索要求として入力させ、入力されたシーン記述言語を検索キーワードとして上記画像検索手段へ渡すことを特徴としている。

【0010】請求項3記載の画像データベース装置は、請求項1記載の装置において、シーン記述言語が、入力画像における各領域の位置を示す位置情報を含むと共に、上記登録手段が、状態遷移手段の処理によって状態が遷移した領域については、上記最終状態に加えて上記位置情報をキーワードとして上記画像記憶手段へ記憶させることを特徴としている。

【0011】請求項4記載の画像データベース装置は、請求項2記載の装置において、上記画像検索手段が検索キーワードに従って画像を検索する際に、検索キーワードが状態遷移モデル中の状態である場合には、該状態の状態遷移モデルにおける階層に応じて、検索キーワードと一致するキーワードを有する画像に得点を付与する一

## 4

方、検索キーワードがシーン記述言語である場合には、画像記憶手段に記憶されている画像のキーワードから遷移ルールを逆行して到達可能な状態遷移モデルの最下位階層の状態を持つシーン記述言語と、上記検索キーワードとしてのシーン記述言語とを比較し、その一致度に応じて上記の画像記憶手段に記憶されている画像に得点を付与すると共に、上記画像提示手段が、所定の得点よりも高い得点を付与された画像を検索結果として提示することを特徴としている。

## 10 【0012】

【作用】請求項1記載の構成によれば、分割された入力画像の各領域から求められた画像特徴量がシーン記述言語に変換され、各領域のシーン記述言語が状態遷移モデルの遷移ルールに適合する限り、各領域の状態が上位概念へ遷移され、状態が遷移した領域についてはその領域の最終状態がキーワードとして、また状態が遷移しなかった領域についてはその領域のシーン記述言語がキーワードとして入力画像と共に画像記憶手段に記憶される。

20 【0013】これにより、状態遷移モデルに適合しない概念を有する画像に対しても、シーン記述言語がキーワードとして自動的に付与されて画像記憶手段に記憶されるので、画像の登録を行う際に登録者がキーワードを付与する必要がないので登録者の作業負担を軽減することが可能となると共に、検索時には、シーン記述言語をキーワードとして与えれば上記画像の検索を行うことができる。すなわち、記憶・検索が可能な画像の種類が画像データベース装置が保持する状態遷移モデルによって限定されないので、画像データベース装置の汎用性が向上すると共に、柔軟な検索が可能となる。

30 【0014】請求項2記載の構成によれば、検索者が入力した検索要求が、画像データベース装置が保持する状態遷移モデル中に状態として存在する場合には、この状態と、上記状態から遷移ルールを逆行して到達可能な下位階層の状態のすべてとが、検索キーワードとして画像検索手段に渡され、画像検索手段がこの検索キーワードと一致するキーワードを有する画像を画像記憶手段から検索し、その検索結果が検索者へ提示される。一方、検索者が入力した検索要求が、画像データベース装置が保持する状態遷移モデル中に状態として存在しない場合には、検索者に検索入力手段により入力させたシーン記述言語が検索キーワードとして画像検索手段へ渡され、画像検索手段は、検索キーワードとして渡されたシーン記述言語に従って画像の検索を行う。

50 【0015】これにより、画像データベース装置が保持している状態遷移モデルに適合しない画像の検索が要求された場合であっても、シーン記述言語を検索要求として入力させ、このシーン記述言語に従って画像の検索を行うことが可能となる。この結果、画像データベース装置で記憶・検索が可能な画像の種類が状態遷移モデルにより限定されることがなくなり、画像データベース装置

の汎用性を向上させることが可能となる。

【0016】請求項3記載の構成によれば、状態遷移手段の処理によって状態が上位階層へ遷移した領域については、その領域の入力画像中の位置が、領域の最終状態と共にキーワードとして入力画像と共に登録されるので、検索を行う際に、検索者が位置情報として例えば「中央部」、領域の状態として例えば「人物」という指定を行うことにより、「中央部に人物が存在する画像」という条件を与えて画像を検索することが可能となる。この結果、検索者が抱えている検索対象の画像イメージをより具体的に画像データベース装置へ指示して検索を行うことが可能となり、検索者の意志が反映された柔軟な検索が可能となると共に、検索効率を向上させることができる。

【0017】請求項4記載の構成によれば、状態遷移モデルにおける状態を検索キーワードとして検索された画像は、上記状態の状態遷移モデルにおける階層に応じて、階層が高くなるほど高い得点が付与され、シーン記述言語を検索キーワードとして検索された画像は、画像記憶手段に記憶されている画像のキーワードから遷移

ルールを逆行して到達可能な状態遷移モデルの最下位階層の状態が有するシーン記述言語と、上記検索キーワードとしてのシーン記述言語とを比較し、その一致度が高くなるほど高い得点が付与され、所定の得点よりも高い得点を付与された画像のみが検索結果として検索者へ提示される。

【0018】これにより、検索者が所望する画像である可能性の高い画像が提示されることとなるため、検索者が検索結果として提示された画像から所望の画像を選択する際等の検索効率を向上させることが可能となると共に、検索者が所望する画像からかけ離れた画像が提示される可能性を減少させることができ、画像データベース装置の信頼性を向上させることができる。

【0019】

【実施例】本発明の一実施例について図1ないし図6に基づいて説明すれば、以下の通りである。本実施例における画像データベース装置は、カラー画像を登録・検索する画像データベース装置であり、画像の登録に関する構成として、図1に示すように、画像とそのキーワードとを蓄積する画像データベース11と、データベースに登録する画像を入力するための画像入力部1（画像入力手段）と、この画像入力部1から入力された画像を蓄積するための入力画像メモリ2と、この入力画像メモリ2に蓄積されたフルカラー画像を入力し、その色数を減色しながら画像全体を複数の領域に分割して限定色画像を作成する限定色画像作成部3（画像分割手段）と、この限定色画像を蓄積する限定色画像メモリ4と、上記の限定色画像をラベリングしてラベリング画像を作成するラベリング部5と、ラベリング画像を蓄積するラベリング画像メモリ6と、各領域の画像特徴量を計算する画像特

徴量計算部7（画像特徴量算出手段）と、画像特徴量をシーン記述言語に変換するシーン記述言語変換部8（シーン記述言語変換手段）と、上記の各領域の状態が遷移するためのルールを蓄積した状態遷移モデルデータベース9（状態遷移モデル記憶手段）と、この状態遷移モデルデータベース9を参照しながら各領域の状態を遷移させる状態遷移部10（状態遷移手段）と、入力画像メモリ2の入力画像をキーワードと共に上記画像データベース11へ登録するデータ登録部12（登録手段）とを備えている。

【0020】ここで、状態遷移モデルについて説明する。状態遷移モデルとは、図6に示すような状態遷移図と、遷移ルールとからなり、同図における例えば“Human”、“Human-hair”等の状態において、予め設定された遷移ルールが満たされれば、矢印で示すように状態が上位階層へ遷移する動作を、満たされる遷移ルールが存在する限り繰り返すことによって、後述するセグメントの状態をより上位概念に近づけるためのモデルである。

【0021】ここで、上記の構成によるカラー画像の登録について、図2に示すフローチャートを参照しながら、以下に説明する。なお、本実施例では、登録されるカラー画像は、赤（R）・緑（G）・青（B）の各要素がそれぞれ8ビット（0から255までの256階調）のデジタルデータで表現されたフルカラー画像であるとする。

【0022】まず、画像データベースに登録しようとするカラー画像が画像入力部1により入力されて入力画像メモリ2に蓄積される（ステップ1、以下、S1のように表記する）。

【0023】次に、この入力画像から物体等を認識するわけであるが、この認識のためには、まず、画像から認識対象となる物体を切り出す必要がある。ここでは、入力画像はフルカラー画像であるため、同じような色によって構成される領域を抽出するカラーセグメンテーションの手法を用いる。つまり、限定色画像作成部3が、入力画像メモリ2から取り出したフルカラーの入力画像から、16色程度に色を限定した限定色画像を作成し（S2）、ラベリング部5が、上記S2で作成された限定色画像において、同一色の連結領域に対して同一のラベルを付与し、同一ラベルを有する領域を一つの連結領域とすることにより画像を複数の領域に分割する（S3）。以後、これらの分割された領域のそれぞれをセグメントと呼ぶこととする。

【0024】次に、画像特徴量計算部7が、各セグメントについて、面積、外接矩形、色、主軸方向、主軸方向に対する占有率、円形度、重心等の画像特徴量をそれぞれ算出し（S4）、シーン記述言語変換部8が、算出された画像特徴量から、各セグメントのシーン記述言語を作成する（S5）。なお、このシーン記述言語は、セグメントの大きさ、入力画像全体を縦横それぞれ3分割し

て9個のブロックに分けた場合に該当セグメントが含まれるブロック位置、セグメントの色、セグメントの形状、およびセグメントの主軸方向をそれぞれ表現する複数種類のパラメータで構成されている。

【0025】続くS6では、状態遷移部10が、各セグメントのシーン記述言語と、状態遷移モデルデータベース9に蓄積されている状態遷移モデルとに基づいて、各セグメントの概念を決定する作業を以下のように行う。

【0026】すなわち、状態遷移部10は、各セグメントに対して上記S5で作成されたシーン記述言語が状態遷移モデルの遷移ルールを満たしているか否かを、状態遷移モデルの下位概念から順に調べ、遷移ルールが満たされていれば、該セグメントの状態を1つ上の階層へ遷移させる。

【0027】より具体的には、図6に示すような状態遷移モデルが与えられ、あるセグメントのシーン記述言語におけるセグメントの色を示すパラメータが「黒」であれば、このセグメントは遷移ルールを満たしていることとなり、最下位の概念である“Color-segment”から1階層上位へ遷移して“Black-segment”となる。さらに、このセグメントのシーン記述言語が、人間の頭髮部分の画像である可能性がある条件を満たしている場合には、さらに遷移ルールが満たされたこととなり、さらに1階層上位へ遷移する。これにより、このセグメントは、人間の頭髮であるという概念“Human-hair”を有するセグメントとして認識される。

【0028】また、上記の状態遷移モデルには、S3における領域分割時の過分割やシェーディング等のために、本来1つの物体であるものが複数セグメントに分割されてしまった物体や、複数色からなる物体等に対応するために、比較的上位の階層において、状態を遷移させずにセグメントどうしの統合のみを行う領域統合ルールとしての遷移ルールも含まれており、この領域統合ルールを満たすセグメントどうしは統合され、新たな1つの概念を有するセグメントとして認識される。

【0029】より具体的には、図6に示すような状態遷移モデルが与えられ、入力画像に含まれるセグメントとして、上記の概念“Human-hair”を有するセグメントと、人間の肌部分の画像であるという概念“Human-skin”を有するセグメントとが存在する場合には、これらのセグメントは領域統合ルールを満たすものとして統合され、人間の顔部分の画像であるという概念“Human-face”を有する1つのセグメントとして認識される。

【0030】上記したS6の動作は、入力画像中のすべてのセグメントの状態が遷移しなくなるまで繰り返され、この動作が終了した後に、データ登録部12が、各セグメントに対して上記S6で求められた最終状態、すなわち例えば上記の“Human-face”等の概念をキーワードとして、上記S1で入力された画像と共に画像データベース11に登録する(S7)。なお、このとき、所定の

階層よりも上位の概念へ遷移したセグメントについては、上記キーワードと共に、シーン記述言語の位置パラメータも画像データベース11へ登録される。また、上記S6において状態が全く遷移しなかったセグメントについては、シーン記述言語がキーワードとして登録される。

【0031】なお、上記した状態遷移モデルでは、状態遷移モデルの下位階層ではシーン記述言語を重要視して状態を遷移させるように遷移ルールが設定されており、上位階層になるに従いシーン記述言語の重要性は低くなり、セグメントどうしの統合を優先的に行うよう領域統合ルールが設定されている。このようにして下位階層ではシーン記述言語により状態を遷移させることにより汎用性を高める一方、上位階層ではシーン記述言語にあまり頼らず、セグメント間の関係等を大局的に見て状態遷移を行うことで、認識の信頼性が高められている。

【0032】なお、ここで、上記S2における限定色画像の作成について、図3に示すフローチャートに基づいてより詳しく説明する。

【0033】まず、限定色画像作成部3は、フルカラーの入力画像のヒストグラムを作成する(S21)。本実施例では、R・G・Bのそれぞれが256階調で表現されているため、画像中の各色は、RGB空間中の一辺256(0から255)の立方体中の点として表される。この一辺256の立方体を、セルと呼ばれる一辺8の立方体に分割する。次に、画像中の画素のRGB値を求め、その画素がRGB空間中でどのセルに属するかを求めて、その画素の属するセルの頻度値を1だけ増加させる。この処理をすべての画素に対して行い、ヒストグラムを作成する。

【0034】次に、作成されたヒストグラム中で、最も頻度値の大きいセルである最頻値セルを検出し(S22)、この最頻値セルの表す色を代表色とし、この最頻値セルを中心として予め与えられている半径の球内に含まれるセルの色を上記代表色で置き換える。このように色が決定されたセルを対象から除外してS22へ戻り、同様の処理を対象セルがなくなるまで繰り返す。これらの処理を行う際の上記半径を適当に調整させることでフルカラー画像の色数を16色程度に減色することができる。つまり、色数が少ない場合には、上記半径を小さくし、多い場合には上記半径を大きくすることにより1つの代表色で表すことのできるセルの数を調整し、色数を希望の数に近づけることができる。

【0035】次に、以上に説明したように画像がキーワードと共に登録されている画像データベース11から、検索者が画像の検索を行うために、本実施例の画像データベース装置が備える構成と、この構成による画像の検索動作について、図4および図5を参照しながら説明する。

【0036】本実施例の画像データベース装置は、前述

した画像の登録に関する構成にさらに加えて、図4に示すように、検索を行う検索者が検索要求を入力するための検索要求入力部21（検索要求入力手段）と、入力された検索要求を解析する検索要求解析部22（検索要求解析手段）と、検索要求の解析結果に基づいて該当する画像を画像データベース11から検索する検索部23（画像検索手段）と、検索部23によって検索された画像を検索者に提示する画像提示部24（画像提示手段）とを備えた構成である。

【0037】上記の構成において、画像データベース11に蓄積されている画像を検索する場合には、まず、検索者は、検索要求入力部21により、検索したい画像を構成している構成物の名称を検索要求として入力する

（S11）。より具体的に説明すると、例えば、検索者が人物を含んだ画像を検索したい場合には、構成物の名称として“Human”と入力することになる。

【0038】検索要求が入力されると、検索要求解析部22が状態遷移モデルデータベース9を探索し、S11で検索要求として入力された構成物の名称と一致する概念が、状態遷移モデルデータベース9に蓄積されている状態遷移モデルに含まれていれば（S12）、S11に戻り、検索要求入力部21が、画像中における上記構成物の位置に関する検索要求を入力するように検索者に対して指示する。

【0039】これにより、検索者は、例えば「真中に人物がいる画像」というような検索要求を作成することが可能となる。この結果、検索者が抱えている検索対象の画像イメージをより具体的に画像データベース装置へ指示して検索を行うことが可能となり、検索者の意志を反映した柔軟な検索が可能となると共に、検索効率が向上するという効果を奏する。ただし、検索者は上記の指示に必ずしも従う必要はなく、位置に関する検索要求を入力せずに次のステップであるS13へ進むこともできる。

【0040】一方、検索要求として入力された構成物の名称と一致する概念が状態遷移モデルデータベース9に蓄積されている状態遷移モデルに存在しなければ（S12）、S11へ戻り、検索要求入力部21が、シーン記述言語による検索要求を検索者に入力させる。なお、シーン記述言語とは、前記したように、対象物の大きさ、入力画像全体に対する位置、色、形状、および主軸方向である。ただし、検索要求としてこれらすべてを入力する必要はない。

【0041】このようにして、状態遷移モデルに存在する概念と必要に応じてその位置、あるいは、所望の概念が状態遷移モデルに存在しない場合はシーン記述言語が、検索要求として検索部23へ渡される。なお、上記S11およびS12は、検索者がすべての検索要求の作成を終了するまで繰り返される（S13）。

【0042】なお、検索者が検索要求を入力するための

検索要求入力部21におけるユーザインターフェイスは、例えばGUI（Graphical User Interface）によるメニュー画面において状態遷移モデルに含まれている概念をアイコンとして表示させ、これらのアイコンから検索者が所望のものを選択すると、選択されたアイコンに対応する概念が検索要求として検索部23へ渡されるようにすることも可能であり、これによれば、検索者にとっての操作性を向上させることができる。また、検索者が所望の構成物に対応する概念がアイコンとして表示されていない場合、すなわち所望の概念が状態遷移モデルに含まれていない場合には、シーン記述言語の各種のパラメータを入力させるよう画面で指示することもできる。

【0043】検索部23は、上記した検索要求に基づいて画像データベース11から該当する画像を検索し（S14）、画像提示部24が検索部23の検索結果を検索者へ提示する（S15）。

【0044】ここで、図6に示す状態遷移モデルの例を用いて、上記の検索動作についてより詳細に説明する。

【0045】検索要求解析部22は、入力された検索要求が、状態遷移モデルに存在する概念であるか、あるいはシーン記述言語であるかによって異なった処理を行う。まず最初に、検索要求として、状態遷移モデルに存在する概念が入力された場合の処理について説明する。ここでは、検索者によって、構成物の名称“Human”が検索要求として入力され、図6に示すように状態遷移モデルに“Human”が概念として存在しているとする。

【0046】検索要求解析部22の解析により、上記の検索要求は状態遷移モデルに存在することが判定すると、検索要求解析部22は、この検索要求“Human”を検索キーワードとして検索部23へ渡し、検索部23はこの検索キーワードに基づいて画像データベース11を探索する。探索の結果、画像データベース11に“Human”がキーワードである画像が登録されていれば、該画像を候補画像とする。さらに、検索要求解析部22は、上記状態遷移モデルにおいて“Human”から1階層だけ下位へ下りた概念である“Human-face”を検索キーワードとして検索部23へ渡し、検索部23はこの検索キーワードと一致するキーワードを有する画像が画像データベース11に登録されているか否かを探索し、存在する場合は該画像を候補画像としてこの“Human-face”の階層に応じた得点を与える。なお、上記の得点は、状態遷移モデルの階層に応じて与えられるものであり、上位の階層に近づくほど高得点が与えられるようになっている。

【0047】上記のような検索キーワードによる探索および得点の付与を、状態遷移モデルの最下位の階層まで順次行ったら、候補画像となった画像の内、あらかじめ設定されている閾値よりも大きな得点を有する画像が、検索結果の画像として画像提示部24によって検索者へ提示される。また、複数種類の検索要求がある場合は、



候補画像となった画像の内、すべての検索要求に対する得点の平均値が上記の閾値よりも大きくなった画像が検索結果として提示される。

【0048】次に、入力された検索要求がシーン記述言語であった場合の処理について説明する。ここでも、図6に示す状態遷移モデルが用いられることとする。

【0049】この場合、検索要求解析部22は、入力されたシーン記述言語を検索キーワードとして検索部23へ渡す。次に、状態遷移モデルデータベース9の状態遷移モデルにおける概念の内、登録画像のキーワードとして画像データベース11に登録されているものが状態遷移モデルから選択される。つまり、例えば図6に示す状態遷移モデルに存在する概念の内、“Human-face”のみがキーワードとして画像と共に画像データベース11に登録されているものとする、この“Human-face”が選択される。

【0050】そして、検索部23が、“Human-face”から下位階層へ、遷移ルールを逆行することにより到達する最下位階層の概念において、検索要求として与えられたシーン記述言語との比較を行う。より詳しくは、状態遷移モデルの下位階層における遷移ルールは、シーン記述言語の各パラメータが満たすべき条件として与えられているものであり、検索要求として与えられたシーン記述言語の各パラメータと、上記の条件としてのシーン記述言語の各パラメータとがどの程度一致しているかに応じて、上記の“Human-face”をキーワードとして画像データベース11に登録されている画像に対して、候補画像としての得点が与えられる。なお、上記の各パラメータの一致度が高いほど、与えられる得点も高くなるものとする。

【0051】上記のように、シーン記述言語の各パラメータに基づく比較および得点の付与を、画像データベース11に登録されているすべてのキーワードに対して行った後、候補画像となった画像の内、あらかじめ設定されている閾値よりも大きな得点を有する画像が、検索結果の画像として画像提示部24によって検索者へ提示される。また、複数種類の検索要求がある場合は、候補画像となった画像の内、すべての検索要求に対する得点の平均値が上記の閾値よりも大きい画像が検索結果として提示される。

【0052】以上のように、本実施例の画像データベース装置は、画像入力部が入力したフルカラーの画像を色に基づいて複数のセグメントに分割し、各セグメントから算出した画像特徴量からシーン記述言語を作成する。さらに、このシーン記述言語に状態遷移モデルを適用することによって各セグメントの状態を可能な限り高レベルの概念に遷移させ、画像データベースに画像を登録する際は、画像と共に、各セグメントの最終状態をキーワードとして登録する。また、この時、上位階層へ遷移したセグメントについては、上記の最終状態と共に画像中

における該セグメントの位置も登録し、上位階層への遷移が行われなかったセグメントについては、シーン記述言語を登録する構成である。

【0053】これにより、登録する画像の各々に対して登録者がキーワードを付与する作業が不要となるので、画像データベースの作成時の登録者の作業負担を大幅に削減することができると共に、登録者の主観によらずに客観的なキーワードが付与されるので、画像データベースの汎用性および信頼性を向上させることができるという効果を奏する。

【0054】また、このようにして登録された画像を検索する際に、状態遷移モデルに存在する概念が検索要求として与えられた場合には、この概念をキーワードとして画像データベースを検索すると共に、上記概念へ到達する遷移ルールにおける下位概念をキーワードとした画像が画像データベースに登録されているか否かの探索を最下位概念まで順次行い、画像が登録されている場合には、その画像のキーワードとしての概念の状態遷移モデルにおける階層に応じた得点が上記画像に与えられ、この得点が所定の閾値よりも大きい画像のみを検索結果として検索者に提示する構成となっている。

【0055】これにより、検索者が与えた検索要求としての概念がキーワードとして登録されていない場合でも、その概念の下位階層の概念をキーワードとして登録されている画像があれば検索結果として提示されるため、検索者の所望する画像が検索結果として提示される可能性を高めることが可能となっている。この結果、より抽象的な高度な概念に基づいた柔軟な検索が可能となるという効果を奏する。また、得点が所定の閾値よりも大きい画像のみが提示されるため、検索者が所望する画像とかけ離れた画像が提示されることを防止して、検索者が所望の画像を選択する負担を軽減することが可能となると共に、画像データベースの信頼性を向上させることができる。

【0056】また、画像の登録の際に、本画像データベース装置が保持する状態遷移モデルにより状態が遷移しなかったセグメントについては、シーン記述言語がキーワードとして画像と共に登録され、検索の際に、本画像データベース装置が保持している状態遷移モデルに存在しない概念が検索要求として検索者から与えられた場合には、シーン記述言語レベルでの検索要求を入力するように検索者に指示を与えると共に、画像データベースにキーワードとして登録されている概念から遷移ルールを逆行することにより到達可能な最下位階層の概念における遷移ルールと、検索者により検索要求として入力されたシーン記述言語との各パラメータを比較し、その一致度に応じて上記キーワードに対応して登録されている画像に得点を与え、この得点が所定の閾値よりも大きい画像のみを検索結果として提示する構成である。

【0057】これにより、本画像データベース装置が保

持っている状態遷移モデルに存在しない概念を有する画像についても、シーン記述言語を利用することにより登録および検索を行うことが可能となるので、画像データベースが状態遷移モデルのみに依存することがなく、画像データベースの汎用性を向上させることができるという効果を奏する。

【0058】なお、本発明は上記実施例に限定されるものではなく、例えば本実施例ではシーン記述言語のパラメータの種類を、セグメントの大きさ、位置、色、形状、および主軸方向として説明したが、データベースに登録する画像の種類に応じてこの他の要素をシーン記述言語とすることも可能である。また、状態遷移モデルについて、登録する画像の種類やデータベースの利用目的に応じて、種々のモデルを適用することができる。

【0059】

【発明の効果】以上のように、本発明の請求項1記載の画像データベース装置は、画像をキーワードと対応づけて記憶する画像記憶手段と、画像を入力する画像入力手段と、入力画像を領域分割する画像分割手段と、分割された各領域の画像特徴量を求める画像特徴量算出手段と、画像特徴量をシーン記述言語に変換するシーン記述言語変換手段と、状態遷移モデルを記憶する状態遷移モデル記憶手段と、各領域の状態を上位階層へ遷移させる処理を、各領域のシーン記述言語が状態遷移モデルの遷移ルールに適合する限り行う状態遷移手段と、上記状態遷移手段の処理によって状態が遷移した領域については該領域の最終状態をキーワードとし、状態が遷移しなかった領域についてはシーン記述言語をキーワードとして、上記入力画像と共に上記画像記憶手段へ記憶させる登録手段とを備えている構成である。

【0060】これにより、状態遷移モデルに適合しない概念を有する画像に対してもシーン記述言語がキーワードとして自動的に付与されるので、登録者の作業負担を軽減することが可能となると共に、記憶・検索が可能な画像の種類が画像データベース装置が保持する状態遷移モデルによって限定されないので、画像データベース装置の汎用性が向上するという効果を奏する。

【0061】請求項2記載の画像データベース装置は、検索者が検索要求を入力する検索要求入力手段と、入力された検索要求を解析して検索を実行する際の検索キーワードを作成する検索要求解析手段と、上記検索キーワードに従って画像記憶手段から画像を検索する画像検索手段と、検索結果の画像を検索者へ提示する画像提示手段とをさらに備えると共に、上記検索要求解析手段が、入力された検索要求が状態遷移モデル中に存在する状態である場合には、該状態および該状態から遷移ルールを逆行して到達可能な状態すべてを検索キーワードとして上記画像検索手段へ渡す一方、入力された検索要求が状態遷移モデル中に状態として存在しない場合には、検索者に検索要求入力手段によりシーン記述言語を検索要求

として入力させ、入力されたシーン記述言語を検索キーワードとして上記画像検索手段へ渡す構成である。

【0062】これにより、画像データベース装置が保持している状態遷移モデルに適合しない画像の検索が要求された場合であっても、シーン記述言語を検索要求として入力させて画像の検索を行うことが可能となる。この結果、画像データベース装置で記憶・検索が可能な画像の種類が状態遷移モデルにより限定されることがなくなり、より柔軟な検索が可能となると共に、画像データベース装置の汎用性を向上させることが可能となるという効果を奏する。

【0063】請求項3記載の画像データベース装置は、シーン記述言語が、入力画像における各領域の位置を示す位置情報を含むと共に、上記登録手段が、状態遷移手段の処理によって状態が遷移した領域については、上記最終状態に加えて上記位置情報をキーワードとして上記画像記憶手段へ記憶させる構成である。

【0064】これにより、画像に含まれる物体をその位置と共に検索要求として入力することが可能となるため、検索者が抱いている検索対象の画像イメージをより具体的に画像データベース装置へ指示して検索を行うことが可能となる。この結果、検索者の意志を反映したより柔軟な検索が可能となり、検索効率を向上させることができるという効果を奏する。

【0065】請求項4記載の画像データベース装置は、上記画像検索手段が検索キーワードに従って画像を検索する際に、検索キーワードが状態遷移モデル中の状態である場合には、該状態の状態遷移モデルにおける階層に応じて、検索キーワードと一致するキーワードを有する画像に得点を付与する一方、検索キーワードがシーン記述言語である場合には、画像記憶手段に記憶されている画像のキーワードから遷移ルールを逆行して到達可能な状態遷移モデルの最下位階層の状態の持つシーン記述言語と、上記検索キーワードとしてのシーン記述言語とを比較し、その一致度に応じて上記の画像記憶手段に記憶されている画像に得点を付与すると共に、上記画像提示手段が、所定の得点よりも高い得点を付与された画像を検索結果として提示する構成である。

【0066】これにより、検索者が所望する画像である可能性の高い画像が提示されるので、検索者が検索結果として提示された画像から所望の画像を選択する際等の検索効率を向上させることが可能となると共に、検索者が所望する画像からかけ離れた画像が提示される可能性が減少するので、検索者が所望の画像を選択する負担を軽減することができると共に、画像データベース装置の信頼性を向上させることができるという効果を奏する。

【図面の簡単な説明】

【図1】本発明の一実施例における画像データベース装置において画像の登録を行うための構成の概略を示すブロック図である。

15

【図 2】上記画像データベース装置における画像の登録処理の流れを示すフローチャートである。

【図 3】上記の画像の登録処理の一部であり、フルカラー画像から限定色画像を作成する処理の流れを示すフローチャートである。

【図 4】上記画像データベース装置において登録されている画像の検索を行うための構成の概略を示すブロック図である。

【図 5】上記画像データベース装置における画像の検索処理の流れを示すフローチャートである。

【図 6】上記画像データベース装置が保持する状態遷移モデルの一例を示す説明図である。

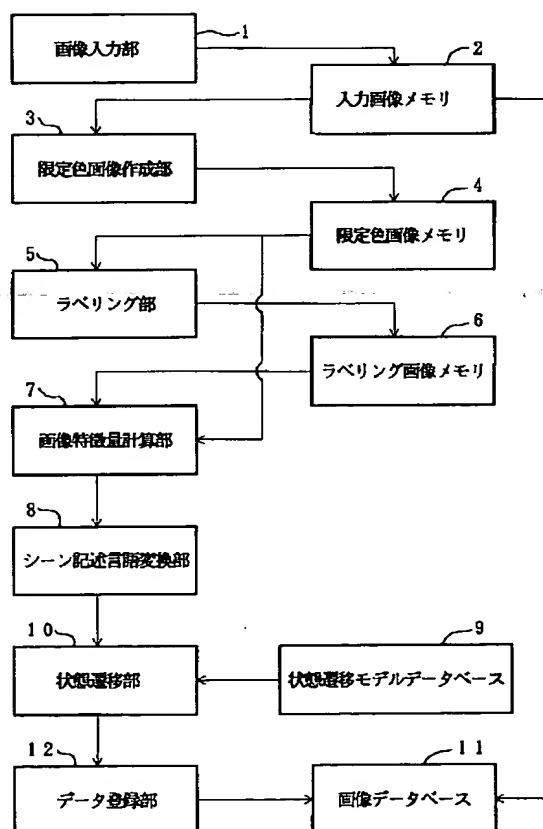
【符号の説明】

1 画像入力部（画像入力手段）

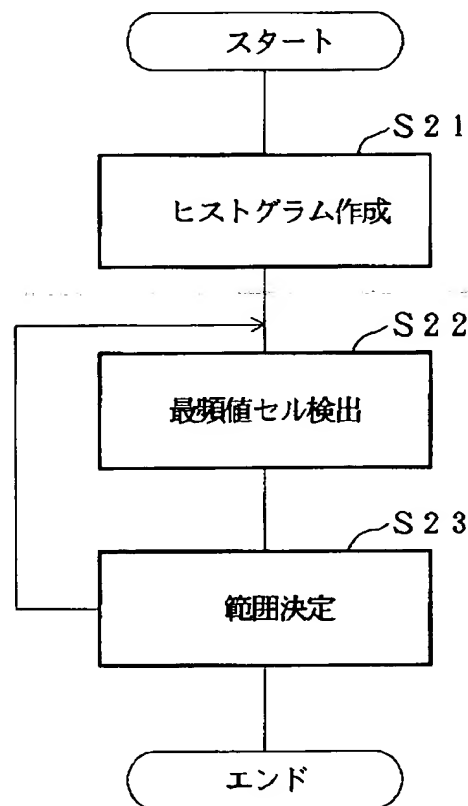
16

- 3 限定色画像作成部（画像分割手段）
- 7 画像特徴量計算部（画像特徴量算出手段）
- 8 シーン記述言語変換部（シーン記述言語変換手段）
- 9 状態遷移モデルデータベース（状態遷移モデル記憶手段）
- 10 状態遷移部（状態遷移手段）
- 11 画像データベース（画像記憶手段）
- 12 データ登録部（登録手段）
- 21 検索要求入力部（検索要求入力手段）
- 22 検索要求解析部（検索要求解析手段）
- 23 検索部（画像検索手段）
- 24 画像提示部（画像提示手段）

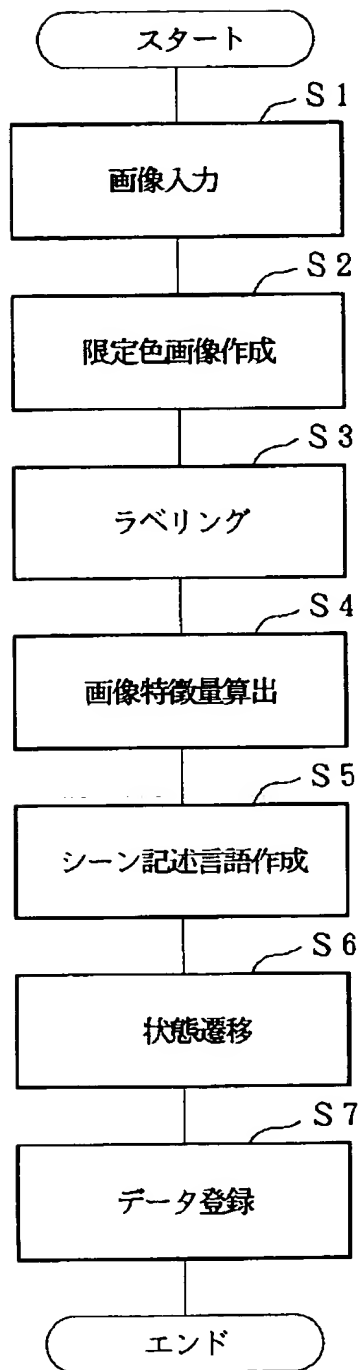
【図 1】



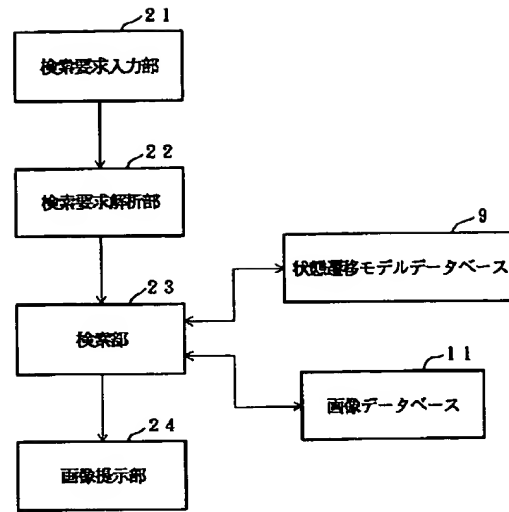
【図 3】



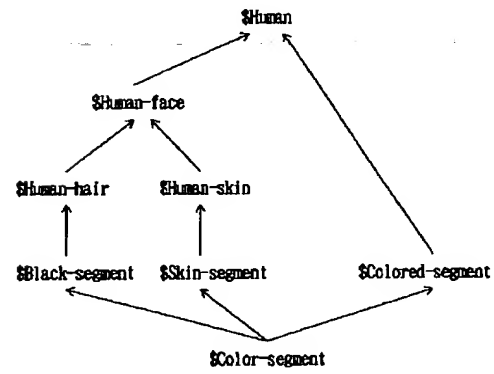
【図 2】



【図 4】



【図 6】



【図 5】

